Letting your fingers do the walking

Jeremy M. Berg

Even before the first three-dimensional structures of DNA binding proteins were determined, it was hoped that simple relationships might exist between the amino acid residues used for sequence-specific nucleic acid recognition and the nucleotide bases that these residues contact. With the elucidation of the amino acid sequences and three-dimensional structures of a large number of DNA binding proteins, it has been discovered that a relatively small number of DNA-binding structural motifs are repeatedly used. Among these, the Cys_2His_2 zinc finger motif—first identified in the transcription factor IIIA (TFIIIA)—has proven to be the most amenable for the elucidation of such recognition rules and for the generation of novel sequence-specific DNA-binding proteins.

Greisman and Pabo[1] have recently reported in *Science* an extension of previous strategies based on phage display that has allowed the identification of arrays of three zinc finger domains that specifically bind to a more diverse set of DNA-binding sites than had previously been demonstrated. With this and other recently developed methods, it may be possible to find zinc finger arrays that will bind to almost any sequence 9–10 nucleotides in length.

The first direct evidence for the mode of interaction between an array of zinc finger domains and its DNA target site came from the cocrystal structure of the three zinc finger domains from Zif268 bound to an oligonucleotide containing its preferred binding sequence[2]. This structure revealed remarkably modular interactions, with each zinc finger domain primarily interacting with a three base pair DNA subsite. These interactions were mediated via four amino acids from each zinc finger domain that fall in positions 1, 2, 3, and 6 relative to the start of an α-helix that is present in each domain. The apparent independence of each domain provided impetus for DNA-binding protein design based on "mixing and matching" zinc finger domains with known subsite specificities[3].

As an alternative approach, zinc finger proteins have been selected using phage display methodologies[4–7]. In these studies, the residues in the recognition helix of one zinc finger domain within a set of three (such as those in Zif268) were randomized, and phage expressing these proteins were selected on the basis of their ability to bind to specific DNA sequences. Such experiments were reasonably successful, but it became clear that the mode of interaction of one zinc finger domain was influenced by neighboring domains, restricting the DNA sequences for which specific zinc finger arrays could be found. Simultaneously randomizing more than one finger would lead to pools of sequences too large to screen. Greisman and Pabo have developed a strategy that allows phage display methods to be effectively utilized, but with the context dependence of the interactions of each zinc finger factored into the selection. The approach is illustrated in Figure 1.

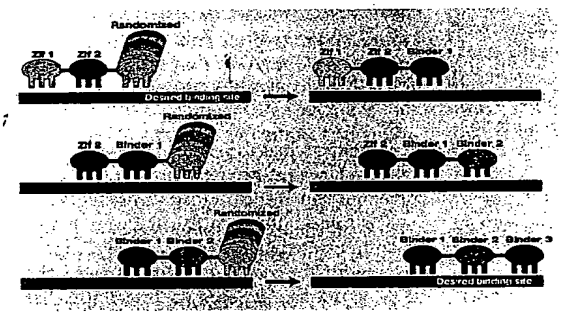

Figure 1. Selection of a three zinc finger protein with desired DNA binding characteristics by "walking" across the binding site.

Proteins comprising three zinc finger domains have sufficient affinity for DNA that they are suitable for selection by phage display methods. Initially, phage were constructed that displayed proteins consisting of two domains from Zif268 (with known DNA-binding preferences) and one domain with a randomized recognition helix. Phage were selected on a hybrid binding site comprising the two subsites for the Zif268-derived domains and a third subsite that was part of the DNA binding site for which a DNA binding protein was sought. After selection of a high-affinity binding protein, a new pool of three domain proteins was constructed with only one Zif268 domain, the new selected domain in the middle, and a new randomized recognition helix domain. This cycle was then repeated, discarding the final Zif268 domain and replacing it with a third new domain to be selected.

The beauty of this approach is that three-domain proteins can be used at all stages, providing sufficiently high binding affinity for effective phage display, but allowing the second and third domains to be selected in the context of neighboring domains. The power of this method was demonstrated by selecting proteins that recognize sites corresponding to a TATA box, a nuclear receptor recognition element, and a p53-binding site. The TATA box result is particularly impressive because this site is nearly devoid of GC base pairs that are crucial for recognition for many of the better characterized zinc finger proteins, including Zif268.

It is gratifying to see that many of the amino acids selected in the canonical DNA recognition residues correspond to those that have been characterized in other contexts. For example, many of the adenine bases appear to be contacted by glutamine or asparagine residues in the selected proteins. Contacts between adenine and the carboxamide groups of these amino acids have recently been observed in the crystal structure of a designed zinc finger protein–DNA complex[8].

The confluence of these studies suggests that our understanding may soon advance to the point that specific binding proteins can be designed without the need for selection. Such tailor-made, zinc finger-based, DNA-binding proteins are likely to find many applications. Addition of other functional modules has already led to the production of novel restriction endonucleases[9,10] and transcription factors[11]. These designed proteins will be powerful tools for research and may soon find applications as therapeutic and diagnostic agents.

Jeremy M. Berg is professor and director of the department of biophysics and biophysical chemistry at the Johns Hopkins University School of Medicine, 725 North Wolfe St., Baltimore, MD 21205 and is a member of the scientific advisory board of Sangamo Biosciences (jeremy_berg@qmail.bs.jhu.edu).

1. Greisman, H.A. and Pabo, C.O. 1997. *Science* 275:657–661.
2. Pavletich, N.P. and Pabo, C.O. 1991. *Science* 252:809–812.
3. Desjarlais, J.R. and Berg, J.M. 1993. *Proc. Natl. Acad. Sci. USA.* 90:2256–2260.
4. Rebar, E.J. and Pabo, C.O. 1994. *Science* 263:671–673.
5. Choo, Y. and Klug, A. 1994. *Proc. Natl. Acad. Sci. USA* 91:11163–11167.
6. Jamieson, A.C., Kim, S.-H., and Wells, J.A. 1994. *Biochemistry* 33:5689–5695.
7. Wu, H., Yang, W.-P., and Barbas, C.F. III. 1995. *Proc. Natl. Acad. Sci. USA.* 92:344–348.
8. Kim, C.A. and Berg, J.M. 1996. *Nature Struct. Biol.* 3:940–945.
9. Kim, Y.-G., Cha, J., and Chandrasegaren, S. 1996. *Proc. Natl. Acad. Sci. USA.* 93:1156–1160.
10. Huang, B., Schaffer, C.J., Li, Q., and Tsai, M.-D. 1996. *J. Pept Chem.* 15:481–489.
11. Choo, Y., Sanchez-Garcia, I., and Klug, A. 1994. *Nature* 272:642–645.